Derivation of Cramér-Rao Bound for Unconditional Maximum Likelihood Direction of Arrival Estimation

Johannes Lindblom

These notes give a detailed proof of the Cramér-Rao bound for UMA:

$$\boldsymbol{B}_u = \frac{\sigma}{2N}\left[\mathrm{Re}\left\{\boldsymbol{D}(\boldsymbol{\theta}_0)^H \boldsymbol{P}^{\perp}_{\boldsymbol{A}(\boldsymbol{\theta}_0)} \boldsymbol{D}(\boldsymbol{\theta}_0) \odot (\boldsymbol{S}\boldsymbol{A}(\boldsymbol{\theta}_0)\boldsymbol{R}(\boldsymbol{\theta}_0)^{-1}\boldsymbol{A}(\boldsymbol{\theta}_0)\boldsymbol{S})\right\}\right]^{-1}. \quad (1)$$

In addition to that, we give show how the concentrated log-likelihood function can be used for deriving the CRB under conditional model assumption. Also, the relation

$$\frac{\boldsymbol{C}_{\text{cml}}}{N} \succeq \frac{\boldsymbol{C}_{\text{uml}}}{N} = \boldsymbol{B}_u$$

is proved.

I. Notation and Useful Properties

The following notation and properties will be given without further explanation.

- $(\cdot)^H$: Hermitian (conjugate) transpose
- $(\cdot)^T$: Transpose
- $(\cdot)^*$: Conjugate
- $\mathrm{tr}\{\cdot\}$: Trace
- $\|\cdot\|_F$: Frobenius norm
- $\boldsymbol{X}^\dagger = (\boldsymbol{X}^H\boldsymbol{X})^{-1}\boldsymbol{X}^H$: Pseudo inverse of the matrix $\boldsymbol{X}$
- $\delta(s,t) = \begin{cases} 1, & s = t \\ 0, & s \neq t \end{cases}$: The Kronecker delta function
- $\boldsymbol{C} = \boldsymbol{A} \odot \boldsymbol{B}$: Hadamard product, $c_{ij} = a_{ij}b_{ij}$
- $\mathbb{E}\{\cdot\}$: Expected value of a random variable.

A. Useful Trace Properties

Assume that $\boldsymbol{B}$ and $\boldsymbol{C}$ are matrices of appropriate dimensions, then $\mathrm{tr}\{\boldsymbol{BC}\} = \mathrm{tr}\{\boldsymbol{CB}\}$. Also, when $\boldsymbol{B}$ and $\boldsymbol{C}$ are square matrices with $\boldsymbol{B}$ Hermitian we have

$$\mathrm{tr}\big\{\boldsymbol{B}(\boldsymbol{C}+\boldsymbol{C}^H)\big\} = \mathrm{tr}\{\boldsymbol{BC}\} + \mathrm{tr}\big\{\boldsymbol{BC}^H\big\} = \mathrm{tr}\{\boldsymbol{BC}\} + \mathrm{tr}\{\boldsymbol{BC}\}^H = 2\,\mathrm{Re}\{\mathrm{tr}\{\boldsymbol{BC}\}\}\,. \quad (2)$$

B. A Matrix Function Derivative

Let $\boldsymbol{X}(\alpha)$ be a matrix function of α. Then

$$\frac{d}{d\alpha}\log\det(\boldsymbol{X}(\alpha)) = \mathrm{tr}\left\{\boldsymbol{X}(\alpha)^{-1}\frac{d}{d\alpha}\boldsymbol{X}(\alpha)\right\} \quad (3)$$

C. Matrix Inversion Lemma

Let $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, and $\boldsymbol{D}$ be matrices of size $m \times m$, $m \times n$, $n \times n$, and $n \times m$, respectively and with $\boldsymbol{A}$ and $\boldsymbol{C}$ invertible. Then

$$(\boldsymbol{A}+\boldsymbol{BCD})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{C}^{-1}+\boldsymbol{DA}^{-1}\boldsymbol{B})^{-1}\boldsymbol{DA}^{-1}.$$

D. Expectation of Four Gaussian Random Variables

Theorem 1. *Let a, b, c, and d be jointly Gaussian distributed. Then*

$$\mathbb{E}\{abcd\} = \mathbb{E}\{ab\}\,\mathbb{E}\{cd\} + \mathbb{E}\{ac\}\,\mathbb{E}\{bd\} + \mathbb{E}\{ad\}\,\mathbb{E}\{bc\} - 2\,\mathbb{E}\{a\}\,\mathbb{E}\{b\}\,\mathbb{E}\{c\}\,\mathbb{E}\{d\}\,.$$

The theorem is proved in [2]. There is the more general scenario of matrix-valued random variables considered.

E. Projection Matrices

With some sloppy notation, we define

$$\boldsymbol{P} \triangleq \boldsymbol{AA}^\dagger = \boldsymbol{A}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H$$

which is the orthogonal projection onto the space spanned by $\boldsymbol{A}$. Since $\boldsymbol{A}$ is a function of $\boldsymbol{\theta}$, $\boldsymbol{P}$ is also a function of $\boldsymbol{\theta}$. We have the following properties:

- $\boldsymbol{P}^\perp = \boldsymbol{I} - \boldsymbol{P}$, $\boldsymbol{PA} = \boldsymbol{A}$, and $\boldsymbol{P}^\perp\boldsymbol{A} = \boldsymbol{0}$

- $\boldsymbol{P}^2 = \boldsymbol{P}$ and $\boldsymbol{P}^H = \boldsymbol{P}$
- Assume that $\boldsymbol{A}$ is a $m \times n$ matrix with full column rank. Then $\boldsymbol{P}$ have n eigenvalues that are 1 and $m - n$ that are 0. Hence, $\operatorname{tr}\{\boldsymbol{P}\} = n$ and $\operatorname{tr}\{\boldsymbol{P}^{\perp}\} = m - n$.

We define $\boldsymbol{A}_k \triangleq \frac{\partial}{\partial \theta_k}\boldsymbol{A}$ and $\boldsymbol{P}_k \triangleq \frac{\partial}{\partial \theta_k}\boldsymbol{P}$. We will use the derivatives in [3, Sec. II-C].

II. Explicit Formula for $\boldsymbol{C}_{\text{UML}}$

Here, we derive the error covariance matrix of the unconditional maximum likelihood estimate $\boldsymbol{C}_{\text{uml}} = \sqrt{N}\,\mathbb{E}\Big\{(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)^H\Big\}$. The UML estimate of the finite-dimensional parameter vector $\boldsymbol{\theta}$ is consistent when $N \to \infty$. Then, it follows that the UML estimate of $\boldsymbol{\theta}$ asymptotically achieves the CRB, $\boldsymbol{B}_u$, see [5]. That is, we get $\boldsymbol{B}_u = \boldsymbol{C}_{\text{uml}}/N$. The derivation of $\boldsymbol{C}_{\text{uml}}$ is given in [1], but here it is more detailed.

We get the UML estimate $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}_0$ as the minimum of

$$V_{\text{uml}}(\boldsymbol{\theta}) = \log\det\left(\boldsymbol{A}\widehat{\boldsymbol{S}}\boldsymbol{A}^H + \widehat{\sigma}\boldsymbol{I}\right). \tag{4}$$

Remember that $\boldsymbol{A}$, $\widehat{\boldsymbol{S}}$, and $\widehat{\sigma}$ are function of $\boldsymbol{\theta}$. From now, we will omit the subscript of $V(\boldsymbol{\theta})$. Since $\hat{\boldsymbol{\theta}}$ minimizes (4), we have

$$V'(\hat{\boldsymbol{\theta}}) = \mathbf{0},$$

where V' is the gradient of V with respect to $\boldsymbol{\theta}$. Also, V'' us defined to be the Hessian of V. A Taylor expansion around $\boldsymbol{\theta}_0$ gives

$$\mathbf{0} = V'(\boldsymbol{\theta}_0) + V''(\boldsymbol{\theta}_0)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) + \cdots. \tag{5}$$

Since $\hat{\boldsymbol{\theta}}$ converges to $\boldsymbol{\theta}_0$ as $N \to \infty$, second- and higher order terms can be neglected. Without affecting the asymptotic properties of $\hat{\boldsymbol{\theta}}$, we can replace $V''(\boldsymbol{\theta}_0)$ with

$$V_0''(\boldsymbol{\theta}_0) = \lim_{N\to\infty} V''(\boldsymbol{\theta}_0),$$

where $V_0''(\boldsymbol{\theta})$ is what we get when $\widehat{\boldsymbol{R}} \to \boldsymbol{R}$ and $\widehat{\sigma} \to \sigma$. Under weak conditions, the inverse of $V_0''(\boldsymbol{\theta})$ exists. Hence, from (5) we get

$$\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) = -\sqrt{N}[V_0''(\boldsymbol{\theta}_0)]^{-1}V'(\boldsymbol{\theta}_0).$$

Then it follows that

$$\boldsymbol{C}_{\text{uml}} = N[V_0''(\boldsymbol{\theta}_0)]^{-1} \lim_{N\to\infty} \mathbb{E}\{V'(\boldsymbol{\theta}_0)V'(\boldsymbol{\theta}_0)^T\} \, [V_0''(\boldsymbol{\theta}_0)]^{-1}.$$

Now, we can determine the gradient, $V'(\boldsymbol{\theta})$, and the Hessian, $V_0''(\boldsymbol{\theta})$.

A. Calculation of Gradient

Differentiation of $V(\boldsymbol{\theta})$ with respect to θ_i gives (3)

$$\frac{\partial}{\partial\theta_i}V(\boldsymbol{\theta}) = \mathrm{tr}\left\{(\boldsymbol{A}\widehat{\boldsymbol{S}}\boldsymbol{A}^H + \widehat{\sigma}\boldsymbol{I})^{-1}\frac{\partial}{\partial\theta_i}(\boldsymbol{P}\widehat{\boldsymbol{S}}\boldsymbol{P})\right\} = \mathrm{tr}\left\{(\boldsymbol{A}\widehat{\boldsymbol{S}}\boldsymbol{A}^H + \widehat{\sigma}\boldsymbol{I})^{-1}\frac{\partial}{\partial\theta_i}(\boldsymbol{P}\widehat{\boldsymbol{R}}\boldsymbol{P} + \widehat{\sigma}\boldsymbol{P}^\perp)\right\}. \tag{6}$$

Using [3, eq. (18)] we get

$$\frac{\partial}{\partial\theta_i}(\boldsymbol{P}\widehat{\boldsymbol{R}}\boldsymbol{P} + \widehat{\sigma}\boldsymbol{P}^\perp) = \boldsymbol{P}\widehat{\boldsymbol{R}}\boldsymbol{P}_i + (\boldsymbol{P}\widehat{\boldsymbol{R}}\boldsymbol{P}_i)^H + \widehat{\sigma}_i\boldsymbol{P}^\perp + \widehat{\sigma}\boldsymbol{P}^\perp\boldsymbol{P}_i^\perp + \left(\widehat{\sigma}\boldsymbol{P}^\perp\boldsymbol{P}_i^\perp\right)^H \tag{7}$$

$$= \boldsymbol{P}\widehat{\boldsymbol{R}}\boldsymbol{P}_i + (\boldsymbol{P}\widehat{\boldsymbol{R}}\boldsymbol{P}_i)^H + \frac{1}{2}\left(\widehat{\sigma}_i\boldsymbol{P}^\perp + \left(\widehat{\sigma}_i\boldsymbol{P}^\perp\right)^H\right) + \widehat{\sigma}\boldsymbol{P}^\perp\boldsymbol{P}_i^\perp + \left(\widehat{\sigma}\boldsymbol{P}^\perp\boldsymbol{P}_i^\perp\right)^H \tag{8}$$

where $\widehat{\sigma}_i \triangleq \frac{\partial}{\partial\theta_i}\widehat{\sigma}$ is real (will be shown later).

Hence, (6) can be written as

$$\frac{\partial}{\partial\theta_i}V(\boldsymbol{\theta}) = 2\,\mathrm{Re}\left\{\mathrm{tr}\left\{(\boldsymbol{A}\widehat{\boldsymbol{S}}\boldsymbol{A}^H + \widehat{\sigma}\boldsymbol{I})^{-1}\left(\boldsymbol{P}\widehat{\boldsymbol{R}}\boldsymbol{P}_i + \widehat{\sigma}\boldsymbol{P}^\perp\boldsymbol{P}_i^\perp + \frac{1}{2}\widehat{\sigma}_i\boldsymbol{P}^\perp\right)\right\}\right\}. \tag{9}$$

Using the matrix inversion lemma, we have

$$(\boldsymbol{A}\widehat{\boldsymbol{S}}\boldsymbol{A}^H + \widehat{\sigma}\boldsymbol{I})^{-1} = \frac{1}{\widehat{\sigma}}\boldsymbol{I} - \frac{1}{\widehat{\sigma}}\boldsymbol{A}(\widehat{\boldsymbol{S}}\boldsymbol{A}^H\boldsymbol{A} + \widehat{\sigma}\boldsymbol{I})^{-1}\widehat{\boldsymbol{S}}\boldsymbol{A}^H$$

and we get

$$\frac{\partial}{\partial\theta_i}V(\boldsymbol{\theta}) = 2\,\mathrm{Re}\left\{\mathrm{tr}\left\{\left[\frac{1}{\widehat{\sigma}}\boldsymbol{I} - \frac{1}{\widehat{\sigma}}\boldsymbol{A}(\widehat{\boldsymbol{S}}\boldsymbol{A}^H\boldsymbol{A} + \widehat{\sigma}\boldsymbol{I})^{-1}\widehat{\boldsymbol{S}}\boldsymbol{A}^H\right]\left[\boldsymbol{P}\widehat{\boldsymbol{R}}\boldsymbol{P}_i + \frac{1}{2}\widehat{\sigma}_i\boldsymbol{P}^\perp + \widehat{\sigma}\boldsymbol{P}^\perp\boldsymbol{P}_i^\perp\right]\right\}\right\}. \tag{10}$$

Using [3, eq. (18)], we observe that

$$\mathrm{tr}\{\boldsymbol{P}^\perp\boldsymbol{P}_i^\perp\} = \mathrm{tr}\left\{\boldsymbol{P}^\perp(-\boldsymbol{A}_i\boldsymbol{A}^\dagger - \boldsymbol{A}\boldsymbol{A}_i^\dagger)\right\} = -\mathrm{tr}\{\boldsymbol{A}_i\boldsymbol{A}^\dagger\boldsymbol{P}^\perp\} - \mathrm{tr}\left\{\boldsymbol{P}^\perp\boldsymbol{A}\boldsymbol{A}_i^\dagger\right\} = 0 \tag{11}$$

By using (2) we get

$$\begin{aligned}
\operatorname{tr}\{\widehat{\sigma}_i \boldsymbol{P}^{\perp}\} &= \widehat{\sigma}_i \operatorname{tr}\{\boldsymbol{P}^{\perp}\} = \widehat{\sigma}_i (m-d) = \operatorname{tr}\Big\{\boldsymbol{P}_i^{\perp} \widehat{\boldsymbol{R}}\Big\} = -\operatorname{tr}\Big\{\boldsymbol{P}_i \widehat{\boldsymbol{R}}\Big\} \\
&= -\operatorname{tr}\Big\{\widehat{\boldsymbol{R}}(\boldsymbol{P}\boldsymbol{P}_i + (\boldsymbol{P}\boldsymbol{P}_i)^H)\Big\} = -2\operatorname{Re}\Big\{\operatorname{tr}\Big\{\widehat{\boldsymbol{R}}\boldsymbol{P}\boldsymbol{P}_i\Big\}\Big\} = -2\operatorname{Re}\Big\{\operatorname{tr}\Big\{\boldsymbol{P}\widehat{\boldsymbol{R}}\boldsymbol{P}_i\Big\}\Big\} \\
&\Leftrightarrow \\
0 &= \frac{1}{2}\operatorname{tr}\{\widehat{\sigma}_i \boldsymbol{P}^{\perp}\} + \operatorname{Re}\Big\{\operatorname{tr}\Big\{\boldsymbol{P}\widehat{\boldsymbol{R}}\boldsymbol{P}_i\Big\}\Big\}. \qquad (12)
\end{aligned}$$

Note that (12) confirms that $\widehat{\sigma}_i$ is real. Since $\widehat{\boldsymbol{S}} = \widehat{\boldsymbol{R}} - \sigma \boldsymbol{I}$ we get

$$\widehat{\boldsymbol{S}}\boldsymbol{A}^H\boldsymbol{A} + \widehat{\sigma}\boldsymbol{I} = (\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H(\widehat{\boldsymbol{R}} - \widehat{\sigma}\boldsymbol{I})\boldsymbol{A}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H\boldsymbol{A} + \widehat{\sigma}\boldsymbol{I} = (\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H\widehat{\boldsymbol{R}}\boldsymbol{A} \qquad (13)$$

By using (12), (13), and $\boldsymbol{P}^{\perp}\boldsymbol{A} = 0$, we write (10) as

$$\begin{aligned}
\frac{\partial}{\partial \theta_i} V(\boldsymbol{\theta}) &= -2\operatorname{Re}\Big\{\operatorname{tr}\Big\{\frac{1}{\widehat{\sigma}}\boldsymbol{A}(\widehat{\boldsymbol{S}}\boldsymbol{A}^H\boldsymbol{A} + \widehat{\sigma}\boldsymbol{I})^{-1}\widehat{\boldsymbol{S}}\boldsymbol{A}^H\boldsymbol{P}\widehat{\boldsymbol{R}}\boldsymbol{P}_i\Big\}\Big\} \\
&= -\frac{2}{\widehat{\sigma}}\operatorname{Re}\Big\{\operatorname{tr}\Big\{\boldsymbol{A}((\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H\widehat{\boldsymbol{R}}\boldsymbol{A})^{-1}\widehat{\boldsymbol{S}}\boldsymbol{A}^H\boldsymbol{P}\widehat{\boldsymbol{R}}\boldsymbol{P}_i\Big\}\Big\} \qquad (14)
\end{aligned}$$

Again, using [3, eq. (18)] gives

$$\boldsymbol{P}_i\boldsymbol{A} = (\boldsymbol{P}^{\perp}\boldsymbol{A}_i\boldsymbol{A}^{\dagger} + (\boldsymbol{P}^{\perp}\boldsymbol{A}_i\boldsymbol{A}^{\dagger})^H)\boldsymbol{A} = \boldsymbol{P}^{\perp}\boldsymbol{A}_i + (\boldsymbol{A}_i\boldsymbol{A}^{\dagger})^H\boldsymbol{P}^{\perp}\boldsymbol{A} = \boldsymbol{P}^{\perp}\boldsymbol{A}_i. \qquad (15)$$

By inserting (15) and [4, eq. (14)] into (14) we get

$$\begin{aligned}
\frac{\partial}{\partial \theta_i} V(\boldsymbol{\theta}) &= -\frac{2}{\widehat{\sigma}}\operatorname{Re}\Big\{\operatorname{tr}\Big\{(\boldsymbol{A}^H\widehat{\boldsymbol{R}}\boldsymbol{A})^{-1}\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{A}^{\dagger}(\widehat{\boldsymbol{S}} - \widehat{\sigma}\boldsymbol{I})\boldsymbol{A}^{\dagger H}\boldsymbol{A}^H\boldsymbol{P}\widehat{\boldsymbol{R}}\boldsymbol{P}^{\perp}\boldsymbol{A}_i\Big\}\Big\} \\
&= -\frac{2}{\widehat{\sigma}}\operatorname{Re}\Big\{\operatorname{tr}\Big\{(\boldsymbol{A}^H\widehat{\boldsymbol{R}}\boldsymbol{A})^{-1}\boldsymbol{A}^H(\widehat{\boldsymbol{R}} - \widehat{\sigma}\boldsymbol{I})\boldsymbol{A}\boldsymbol{A}^{\dagger}\widehat{\boldsymbol{R}}\boldsymbol{P}^{\perp}\boldsymbol{A}_i\Big\}\Big\} \\
&= -\frac{2}{\widehat{\sigma}}\operatorname{Re}\Big\{\operatorname{tr}\Big\{(\boldsymbol{A}^H\widehat{\boldsymbol{R}}\boldsymbol{A})^{-1}(\boldsymbol{A}^H\widehat{\boldsymbol{R}}\boldsymbol{A} - \widehat{\sigma}\boldsymbol{A}^H\boldsymbol{A})\boldsymbol{A}^{\dagger}\widehat{\boldsymbol{R}}\boldsymbol{P}^{\perp}\boldsymbol{A}_i\Big\}\Big\} \\
&= -\frac{2}{\widehat{\sigma}}\operatorname{Re}\operatorname{tr}\Big\{((\boldsymbol{A}^H\boldsymbol{A})^{-1} - \widehat{\sigma}(\boldsymbol{A}^H\widehat{\boldsymbol{R}}\boldsymbol{A})^{-1})\boldsymbol{A}^H\widehat{\boldsymbol{R}}\boldsymbol{P}^{\perp}\boldsymbol{A}_i\Big\} \\
&= 2\operatorname{Re}\Big\{\operatorname{tr}\Big\{\Big((\boldsymbol{A}^H\widehat{\boldsymbol{R}}\boldsymbol{A})^{-1} - \frac{1}{\widehat{\sigma}}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\Big)\boldsymbol{A}^H\widehat{\boldsymbol{R}}\boldsymbol{P}^{\perp}\boldsymbol{A}_i\Big\}\Big\}. \qquad (16)
\end{aligned}$$

We note that the only non-zero column of $\boldsymbol{A}_i$ is $\boldsymbol{d}_i$. Define $\widehat{\boldsymbol{\rho}}_i{}^H$ to be the i:th row of

$$\Big((\boldsymbol{A}^H\widehat{\boldsymbol{R}}\boldsymbol{A})^{-1} - \frac{1}{\widehat{\sigma}}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\Big)\boldsymbol{A}^H.$$

Hence, we can write (16) as

$$\frac{\partial}{\partial\theta_i}V(\boldsymbol{\theta}) = 2\,\mathrm{Re}\Big\{\widehat{\boldsymbol{\rho}}_i{}^H\widehat{\boldsymbol{R}}\boldsymbol{P}^{\perp}\boldsymbol{d}_i\Big\}\,. \tag{17}$$

When $N \to \infty$ we can replace $\widehat{\sigma}$ and $\widehat{\boldsymbol{R}}$ with σ and $\boldsymbol{R}$ respectively. This follows from the following reasoning: We write $\widehat{\boldsymbol{R}} = \boldsymbol{R} + \widetilde{\boldsymbol{R}} = \boldsymbol{A}^H\boldsymbol{S}\boldsymbol{A} + \sigma\boldsymbol{I} + \widetilde{\boldsymbol{R}}$. Then we have $\boldsymbol{A}^H\widehat{\boldsymbol{R}}\boldsymbol{P}^{\perp} = \boldsymbol{A}^H\widetilde{\boldsymbol{R}}\boldsymbol{P}^{\perp} \to \boldsymbol{0}$, $N \to \infty$. For the scalar case we have

$$\frac{1}{r+\tilde{r}} = \frac{1/r}{1+\tilde{r}/r} = \frac{1}{r}\left(1 - \frac{\tilde{r}}{r} + \dots\right) \approx \frac{1}{r} - \frac{\tilde{r}}{r^2}.$$

The term $\frac{\tilde{r}}{r^2} \to 0$ as $\tilde{r} \to 0$. This can be generalized to the inverse of a matrix. Hence we get

$$\begin{aligned}\frac{\partial}{\partial\theta_i}V(\boldsymbol{\theta}) &= 2\,\mathrm{Re}\Big\{\mathrm{tr}\Big\{\Big((\boldsymbol{A}^H\widehat{\boldsymbol{R}}\boldsymbol{A})^{-1} - \frac{1}{\widehat{\sigma}}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\Big)\,\boldsymbol{A}^H\widehat{\boldsymbol{R}}\boldsymbol{P}^{\perp}\boldsymbol{A}_i\Big\}\Big\}\\ &= 2\,\mathrm{Re}\Big\{\mathrm{tr}\Big\{\Big((\boldsymbol{A}^H\boldsymbol{R}\boldsymbol{A})^{-1} - \frac{1}{\sigma}(\boldsymbol{A}^H\boldsymbol{A})^{-1} + \text{ higher order terms }\Big)\,\boldsymbol{A}^H\widetilde{\boldsymbol{R}}\boldsymbol{P}^{\perp}\boldsymbol{A}_i\Big\}\Big\}\end{aligned}$$

The higher order terms depend linearly on $\widetilde{\boldsymbol{R}}$. Therefore, these terms can be neglected since $\widetilde{\boldsymbol{R}} \to \boldsymbol{0}$. Due to the projection matrix, we can replace $\widetilde{\boldsymbol{R}}$ with $\widehat{\boldsymbol{R}}$. Hence, we have

$$\boldsymbol{\rho}_i{}^H \text{ is the } i\text{:th row of } \Big((\boldsymbol{A}^H\boldsymbol{R}\boldsymbol{A})^{-1} - \frac{1}{\sigma}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\Big)\,\boldsymbol{A}^H. \tag{18}$$

B. Calculation of the Hessian $V_0''(\boldsymbol{\theta})$

First, we determine the Hessian of $V(\boldsymbol{\theta})$. Second, we let $N \to \infty$.

$$\begin{aligned}[V''(\boldsymbol{\theta})]_{ij} =& 2\,\mathrm{Re}\Big\{\mathrm{tr}\Big\{\frac{\partial}{\partial\theta_j}\Big[(\boldsymbol{A}^H\widehat{\boldsymbol{R}}\boldsymbol{A})^{-1} - \frac{1}{\widehat{\sigma}}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\Big]\,\boldsymbol{A}^H\widehat{\boldsymbol{R}}\boldsymbol{P}^{\perp}\boldsymbol{A}_i\Big\}\Big\}\\ &+ 2\,\mathrm{Re}\Big\{\mathrm{tr}\Big\{\Big[(\boldsymbol{A}^H\widehat{\boldsymbol{R}}\boldsymbol{A})^{-1} - \frac{1}{\widehat{\sigma}}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\Big]\,\Big[\boldsymbol{A}_j{}^H\widehat{\boldsymbol{R}}\boldsymbol{P}^{\perp}\boldsymbol{A}_i + \boldsymbol{A}^H\widehat{\boldsymbol{R}}\boldsymbol{P}_j^{\perp}\boldsymbol{A}_i + \boldsymbol{A}^H\widehat{\boldsymbol{R}}\boldsymbol{P}^{\perp}\boldsymbol{A}_{ij}\Big]\Big\}\Big\}.\end{aligned} \tag{19}$$

When $N \to \infty$, $\widehat{\boldsymbol{R}} \to \boldsymbol{R}$. Since $\boldsymbol{R} = \boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^H + \sigma\boldsymbol{I}$, we have

$$\begin{aligned}\boldsymbol{A}^H\boldsymbol{R}\boldsymbol{P}^{\perp} &= (\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^H + \sigma\boldsymbol{A}^H)\boldsymbol{P}^{\perp} = \boldsymbol{0},\\ \boldsymbol{A}_j{}^H\boldsymbol{R}\boldsymbol{P}^{\perp} &= (\boldsymbol{A}_j{}^H\boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^H + \sigma\boldsymbol{A}_j)\boldsymbol{P}^{\perp} = \sigma\boldsymbol{A}_j{}^H\boldsymbol{P}^{\perp}.\end{aligned}$$

Hence, we get

$$[V_0''(\boldsymbol{\theta})]_{ij} = 2\,\mathrm{Re}\left\{\mathrm{tr}\left\{\left[(\boldsymbol{A}^H\boldsymbol{R}\boldsymbol{A})^{-1} - \frac{1}{\sigma}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\right]\left[\sigma\boldsymbol{A}_j{}^H\boldsymbol{P}^{\perp} + \boldsymbol{A}^H\boldsymbol{R}\boldsymbol{P}_j^{\perp}\right]\boldsymbol{A}_i\right\}\right\}. \tag{20}$$

Moreover, using [3, eq. (18)] we have

$$\begin{aligned}\boldsymbol{A}^H\boldsymbol{R}\boldsymbol{P}_j^{\perp} &= -\boldsymbol{A}^H\boldsymbol{R}\boldsymbol{P}_j = -\boldsymbol{A}^H\boldsymbol{R}(\boldsymbol{P}_j\boldsymbol{P} + \boldsymbol{P}\boldsymbol{P}_j)\\ &= -\boldsymbol{A}^H\boldsymbol{R}(\boldsymbol{P}\boldsymbol{A}_j\boldsymbol{A}^{\dagger} + (\boldsymbol{P}\boldsymbol{A}_j\boldsymbol{A}^{\dagger})^H)\boldsymbol{P}^{\perp} = -\boldsymbol{A}^H\boldsymbol{R}\boldsymbol{A}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}_j{}^H\boldsymbol{P}^{\perp}.\end{aligned} \tag{21}$$

Inserting (21) into (20) yields

$$[V_0''(\boldsymbol{\theta})]_{ij} = 2\,\mathrm{Re}\left\{\mathrm{tr}\left\{\left[(\boldsymbol{A}^H\boldsymbol{R}\boldsymbol{A})^{-1} - \frac{1}{\sigma}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\right]\left[\sigma\boldsymbol{I} - \boldsymbol{A}^H\boldsymbol{R}\boldsymbol{A}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\right]\boldsymbol{A}_j{}^H\boldsymbol{P}^{\perp}\boldsymbol{A}_i\right\}\right\}. \tag{22}$$

Note that the only non-zero element in $\boldsymbol{A}_j{}^H\boldsymbol{P}^{\perp}\boldsymbol{A}_i$ is the j, i:th. Therefore,

$$[V_0''(\boldsymbol{\theta})]_{ij} = 2\,\mathrm{Re}\left\{\left[\left[(\boldsymbol{A}^H\boldsymbol{R}\boldsymbol{A})^{-1} - \frac{1}{\sigma}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\right]\left[\sigma\boldsymbol{I} - \boldsymbol{A}^H\boldsymbol{R}\boldsymbol{A}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\right]\right]_{i,j}\right\}\boldsymbol{d}_j{}^H\boldsymbol{P}^{\perp}\boldsymbol{d}_i, \tag{23}$$

where

$$\begin{aligned}&\left[(\boldsymbol{A}^H\boldsymbol{R}\boldsymbol{A})^{-1} - \frac{1}{\sigma}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\right]\left[\sigma\boldsymbol{I} - \boldsymbol{A}^H\boldsymbol{R}\boldsymbol{A}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\right]\\ &= \sigma(\boldsymbol{A}^H\boldsymbol{R}\boldsymbol{A})^{-1} - 2(\boldsymbol{A}^H\boldsymbol{A})^{-1} + \frac{1}{\sigma}(\boldsymbol{A}^H\boldsymbol{A})^{-1}(\boldsymbol{A}^H\boldsymbol{R}\boldsymbol{A})(\boldsymbol{A}^H\boldsymbol{A})^{-1}\\ &= \sigma(\boldsymbol{A}^H(\boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^H + \sigma\boldsymbol{I})\boldsymbol{A})^{-1} - 2(\boldsymbol{A}^H\boldsymbol{A})^{-1} + \frac{1}{\sigma}(\boldsymbol{A}^H\boldsymbol{A})^{-1}(\boldsymbol{A}^H(\boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^H + \sigma\boldsymbol{I})\boldsymbol{A})(\boldsymbol{A}^H\boldsymbol{A})^{-1}\\ &= \sigma(\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^H\boldsymbol{A} + \sigma\boldsymbol{A}^H\boldsymbol{A})^{-1} - 2(\boldsymbol{A}^H\boldsymbol{A})^{-1} + \frac{1}{\sigma}(\boldsymbol{S} + \sigma(\boldsymbol{A}^H\boldsymbol{A})^{-1})\\ &= \sigma(\boldsymbol{S}\boldsymbol{A}^H\boldsymbol{A} + \sigma\boldsymbol{I})^{-1}(\boldsymbol{A}^H\boldsymbol{A})^{-1} - (\boldsymbol{A}^H\boldsymbol{A})^{-1} + \frac{\boldsymbol{S}}{\sigma} = \{\text{Use matrix inversion lemma}\}\\ &= \sigma\left(\frac{\boldsymbol{I}}{\sigma} - \frac{\boldsymbol{S}}{\sigma}\left(\boldsymbol{I} + \frac{1}{\sigma}\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{S}\right)^{-1}\frac{\boldsymbol{A}^H\boldsymbol{A}}{\sigma}\right)(\boldsymbol{A}^H\boldsymbol{A})^{-1} - (\boldsymbol{A}^H\boldsymbol{A})^{-1} + \frac{\boldsymbol{S}}{\sigma}\\ &= \frac{\boldsymbol{S}}{\sigma}\left(\boldsymbol{I} - \left(\boldsymbol{I} + \frac{1}{\sigma}\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{S}\right)^{-1}\right) = \{\text{Use matrix inversion lemma}\}\\ &= \frac{\boldsymbol{S}}{\sigma}\left(\boldsymbol{I} + \frac{1}{\sigma}\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{S}\right)^{-1}\frac{1}{\sigma}\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{S} = \frac{\boldsymbol{S}}{\sigma}(\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{S} + \sigma\boldsymbol{I})^{-1}\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{S}.\end{aligned}$$

We note that

$$(\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{S}+\sigma\boldsymbol{I})^{-1}\boldsymbol{A}^H\boldsymbol{R} = (\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{S}+\sigma\boldsymbol{I})^{-1}\boldsymbol{A}^H(\boldsymbol{A}\boldsymbol{S}\boldsymbol{A}+\sigma\boldsymbol{I})$$

$$= (\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{S}+\sigma\boldsymbol{I})^{-1}(\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{S}+\sigma\boldsymbol{I})\boldsymbol{A}^H = \boldsymbol{A}^H$$

$$\Leftrightarrow$$

$$(\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{S}+\sigma\boldsymbol{I})^{-1}\boldsymbol{A}^H = \boldsymbol{A}^H\boldsymbol{R}^{-1}. \tag{24}$$

Now, we can write (23) as

$$[V_0''(\boldsymbol{\theta})]_{ij} = \frac{2}{\sigma}\operatorname{Re}\Big\{\big[\boldsymbol{S}\boldsymbol{A}^H\boldsymbol{R}^{-1}\boldsymbol{A}\boldsymbol{S}\big]_{i,j}\Big\}\,\boldsymbol{d}_j{}^H\boldsymbol{P}^\perp\boldsymbol{d}_i, \tag{25}$$

which finally gives us the Hessian

$$V_0''(\boldsymbol{\theta}) = \frac{2}{\sigma}\operatorname{Re}\big\{\boldsymbol{D}^H\boldsymbol{P}^\perp\boldsymbol{D}\odot(\boldsymbol{S}\boldsymbol{A}^H\boldsymbol{R}^{-1}\boldsymbol{A}\boldsymbol{S})^T\big\} \tag{26}$$

III. Calculation of Error Covariance

Since the $V'(\boldsymbol{\theta})$ is linear in $\widehat{\boldsymbol{R}}$, the second order moments of $V'(\boldsymbol{\theta})$ are of the form $\mathbb{E}\{abcd\}$. We define

$$\begin{aligned}[\boldsymbol{C}]_{ij} &\triangleq \lim_{N\to\infty}\mathbb{E}\big\{V'(\boldsymbol{\theta}_0)V'(\boldsymbol{\theta}_0)^T\big\}\\ &= 4\lim_{N\to\infty}\mathbb{E}\Big\{\operatorname{Re}\Big\{\boldsymbol{\rho}_i{}^H\widehat{\boldsymbol{R}}\boldsymbol{P}^\perp\boldsymbol{d}_i\Big\}\operatorname{Re}\Big\{\boldsymbol{\rho}_j{}^H\widehat{\boldsymbol{R}}\boldsymbol{P}^\perp\boldsymbol{d}_j\Big\}\Big\}\\ &= 2\lim_{N\to\infty}\mathbb{E}\Big\{\operatorname{Re}\Big\{\boldsymbol{\rho}_i{}^H\widehat{\boldsymbol{R}}\boldsymbol{P}^\perp\boldsymbol{d}_i\boldsymbol{\rho}_j{}^H\widehat{\boldsymbol{R}}\boldsymbol{P}^\perp\boldsymbol{d}_j\Big\}+\operatorname{Re}\Big\{\boldsymbol{\rho}_i{}^H\widehat{\boldsymbol{R}}\boldsymbol{P}^\perp\boldsymbol{d}_i\boldsymbol{d}_j{}^H\widehat{\boldsymbol{R}}\boldsymbol{P}^\perp\boldsymbol{\rho}_j\Big\}\Big\}.\end{aligned}$$

Note that

$$\widehat{\boldsymbol{R}}\boldsymbol{P}^\perp = \frac{1}{N}\sum_{t=1}^N\boldsymbol{y}(t)\boldsymbol{y}^H(t)\boldsymbol{P}^\perp = \frac{1}{N}\sum_{t=1}^N\boldsymbol{y}(t)(\boldsymbol{x}(t)^H\boldsymbol{A}^H+\boldsymbol{w}^H(t))\boldsymbol{P}^\perp = \frac{1}{N}\sum_{t=1}^N\boldsymbol{y}(t)\boldsymbol{w}^H(t)$$

and $\mathbb{E}\big\{\boldsymbol{y}(t)\boldsymbol{w}^H(t)\big\} = \mathbb{E}\big\{(\boldsymbol{A}\boldsymbol{x}(t)+\boldsymbol{w}(t))\boldsymbol{w}^H(t)\big\} = \sigma\boldsymbol{I}$ and $\boldsymbol{\rho}_i\boldsymbol{P}^\perp = \boldsymbol{0}$. So, by using Theorem 1, (identifying $a = \boldsymbol{\rho}_i{}^H\boldsymbol{y}(t)$, $b = \boldsymbol{w}^H(t)\boldsymbol{P}^\perp\boldsymbol{d}_i$, $c = \boldsymbol{\rho}_j{}^H\boldsymbol{y}(s)$, and $d = \boldsymbol{w}^H(s)\boldsymbol{P}^\perp\boldsymbol{d}_j$) we get

$$
\begin{aligned}
\mathbb{E}\Big\{\boldsymbol{\rho}_i{}^H\widehat{\boldsymbol{R}}\boldsymbol{P}^{\perp}\boldsymbol{d}_i\boldsymbol{\rho}_j{}^H\widehat{\boldsymbol{R}}\boldsymbol{P}^{\perp}\boldsymbol{d}_j\Big\} =& \frac{1}{N^2}\sum_{t=1}^{N}\sum_{s=1}^{N}\mathbb{E}\{\boldsymbol{\rho}_i{}^H\boldsymbol{y}(t)\boldsymbol{w}^H(t)\boldsymbol{P}^{\perp}\boldsymbol{d}_i\boldsymbol{\rho}_j{}^H\boldsymbol{y}(s)\boldsymbol{w}^H(s)\boldsymbol{P}^{\perp}\boldsymbol{d}_j\}\\
=& \frac{1}{N^2}\sum_{t=1}^{N}\sum_{s=1}^{N}\big(\mathbb{E}\{\boldsymbol{\rho}_i{}^H\boldsymbol{y}(t)\boldsymbol{w}^H(t)\boldsymbol{P}^{\perp}\boldsymbol{d}_i\}\,\mathbb{E}\{\boldsymbol{\rho}_j{}^H\boldsymbol{y}(s)\boldsymbol{w}^H(s)\boldsymbol{P}^{\perp}\boldsymbol{d}_j\}\\
&+\mathbb{E}\big\{\boldsymbol{\rho}_i{}^H\boldsymbol{y}(t)\boldsymbol{\rho}_j{}^H\boldsymbol{y}(s)\big\}\,\mathbb{E}\big\{\boldsymbol{w}^H(t)\boldsymbol{P}^{\perp}\boldsymbol{d}_i\boldsymbol{w}^H(s)\boldsymbol{P}^{\perp}\boldsymbol{d}_j\big\}\\
&+\mathbb{E}\big\{\boldsymbol{\rho}_i{}^H\boldsymbol{y}(t)\boldsymbol{w}^H(s)\boldsymbol{P}^{\perp}\boldsymbol{d}_j\big\}\,\mathbb{E}\big\{\boldsymbol{w}^H(t)\boldsymbol{P}^{\perp}\boldsymbol{d}_i\boldsymbol{\rho}_j{}^H\boldsymbol{y}(s)\big\}\\
&-2\,\mathbb{E}\big\{\boldsymbol{\rho}_i{}^H\boldsymbol{y}(t)\big\}\,\mathbb{E}\big\{\boldsymbol{w}^H(t)\boldsymbol{P}^{\perp}\boldsymbol{d}_i\big\}\,\mathbb{E}\big\{\boldsymbol{\rho}_j{}^H\boldsymbol{y}(s)\big\}\,\mathbb{E}\big\{\boldsymbol{w}^H(s)\boldsymbol{P}^{\perp}\boldsymbol{d}_j\big\}\big)\\
=& \frac{1}{N^2}\sum_{t=1}^{N}\sum_{s=1}^{N}\big(\sigma\boldsymbol{\rho}_i{}^H\boldsymbol{P}^{\perp}\boldsymbol{d}_i\sigma\boldsymbol{\rho}_j\boldsymbol{P}^{\perp}\boldsymbol{d}_j\\
&+\mathbb{E}\big\{\boldsymbol{\rho}_i{}^H\boldsymbol{y}(t)\boldsymbol{y}^T(s)\boldsymbol{\rho}_j^*\big\}\,\mathbb{E}\big\{\boldsymbol{d}_i{}^T\boldsymbol{P}^{\perp}\boldsymbol{w}^*(t)\boldsymbol{w}^H(s)\boldsymbol{P}^{\perp}\boldsymbol{d}_j\big\}\\
&+\sigma\delta(t,s)\boldsymbol{\rho}_i{}^H\boldsymbol{P}^{\perp}\boldsymbol{d}_j\,\mathbb{E}\big\{\boldsymbol{w}^H(t)\boldsymbol{P}^{\perp}\boldsymbol{d}_i\boldsymbol{\rho}_j{}^H\boldsymbol{y}(s)\big\}\\
&-2\,\mathbb{E}\big\{\boldsymbol{\rho}_i{}^H\boldsymbol{y}(t)\big\}\,\mathbb{E}\big\{\boldsymbol{w}^H(t)\boldsymbol{P}^{\perp}\boldsymbol{d}_i\big\}\,\mathbb{E}\big\{\boldsymbol{\rho}_j{}^H\boldsymbol{y}(s)\big\}\,\mathbb{E}\big\{\boldsymbol{w}^H(s)\boldsymbol{P}^{\perp}\boldsymbol{d}_j\big\}\big)\\
&=0
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{E}\Big\{\boldsymbol{\rho}_i{}^H\widehat{\boldsymbol{R}}\boldsymbol{P}^{\perp}\boldsymbol{d}_i\boldsymbol{d}_j{}^H\boldsymbol{P}^{\perp}\widehat{\boldsymbol{R}}\boldsymbol{\rho}_j\Big\} =& \frac{1}{N^2}\sum_{t=1}^{N}\sum_{s=1}^{N}\mathbb{E}\{\boldsymbol{\rho}_i{}^H\boldsymbol{y}(t)\boldsymbol{w}^H(t)\boldsymbol{P}^{\perp}\boldsymbol{d}_i\boldsymbol{d}_j{}^H\boldsymbol{P}^{\perp}\boldsymbol{w}(s)\boldsymbol{y}^H(s)\boldsymbol{\rho}_j\}\\
=& \frac{1}{N^2}\sum_{t=1}^{N}\sum_{s=1}^{N}\big(\mathbb{E}\{\boldsymbol{\rho}_i{}^H\boldsymbol{y}(t)\boldsymbol{w}^H(t)\boldsymbol{P}^{\perp}\boldsymbol{d}_i\}\,\mathbb{E}\{\boldsymbol{d}_j{}^H\boldsymbol{P}^{\perp}\boldsymbol{w}(s)\boldsymbol{y}^H(s)\boldsymbol{\rho}_j\}\\
&+\mathbb{E}\big\{\boldsymbol{\rho}_i{}^H\boldsymbol{y}(t)\boldsymbol{d}_j{}^H\boldsymbol{P}^{\perp}\boldsymbol{w}(s)\big\}\,\mathbb{E}\big\{\boldsymbol{w}^H(t)\boldsymbol{P}^{\perp}\boldsymbol{d}_i\boldsymbol{y}^H(s)\boldsymbol{\rho}_j\big\}\\
&+\mathbb{E}\big\{\boldsymbol{\rho}_i{}^H\boldsymbol{y}(t)\boldsymbol{y}^H(s)\boldsymbol{\rho}_j\big\}\,\mathbb{E}\big\{\boldsymbol{w}^H(t)\boldsymbol{P}^{\perp}\boldsymbol{d}_i\boldsymbol{d}_j{}^H\boldsymbol{P}^{\perp}\boldsymbol{w}(s)\big\}\\
&-2\,\mathbb{E}\big\{\boldsymbol{\rho}_i{}^H\boldsymbol{y}(t)\big\}\,\mathbb{E}\big\{\boldsymbol{w}^H(t)\boldsymbol{P}^{\perp}\boldsymbol{d}_i\big\}\,\mathbb{E}\big\{\boldsymbol{d}_j{}^H\boldsymbol{P}^{\perp}\boldsymbol{w}(s)\big\}\,\mathbb{E}\big\{\boldsymbol{y}^H(s)\boldsymbol{\rho}_j\big\}\big)\\
=& \frac{1}{N^2}\sum_{t=1}^{N}\sum_{s=1}^{N}\big(\mathbb{E}\big\{\boldsymbol{\rho}_i{}^H\boldsymbol{y}(t)\boldsymbol{d}_j{}^H\boldsymbol{P}^{\perp}\boldsymbol{w}(s)\big\}\,\mathbb{E}\big\{\boldsymbol{\rho}^T{}_j\boldsymbol{y}^*(s)\boldsymbol{w}^H(t)\boldsymbol{P}^{\perp}\boldsymbol{d}_i\big\}\\
&+\boldsymbol{\rho}_i{}^H\boldsymbol{R}\delta(t,s)\boldsymbol{\rho}_j\boldsymbol{d}_j{}^H\boldsymbol{P}^{\perp}\sigma\boldsymbol{I}\delta(t,s)\boldsymbol{P}^{\perp}\boldsymbol{d}_i\big)\\
=& \frac{\sigma}{N}\boldsymbol{\rho}_i{}^H\boldsymbol{R}\boldsymbol{\rho}_j\boldsymbol{d}_j{}^H\boldsymbol{P}^{\perp}\boldsymbol{d}_i
\end{aligned}
$$

Note that $\boldsymbol{\rho}_i{}^H \boldsymbol{R} \boldsymbol{\rho}_j$ is element (i, j) of

$$\begin{aligned}
&\left[(\boldsymbol{A}^H\boldsymbol{R}\boldsymbol{A})^{-1} - \frac{1}{\sigma}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\right] \boldsymbol{A}^H\boldsymbol{R}\boldsymbol{A} \left[(\boldsymbol{A}^H\boldsymbol{R}\boldsymbol{A})^{-1} - \frac{1}{\sigma}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\right] \\
&= \left[(\boldsymbol{A}^H(\boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^H + \sigma\boldsymbol{I})\boldsymbol{A})^{-1} - \frac{1}{\sigma}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\right] \boldsymbol{A}^H\boldsymbol{R}\boldsymbol{A} \left[(\boldsymbol{A}^H(\boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^H + \sigma\boldsymbol{I})\boldsymbol{A})^{-1} - \frac{1}{\sigma}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\right] \\
&= \left[(\boldsymbol{S}\boldsymbol{A}^H\boldsymbol{A} + \sigma\boldsymbol{I})^{-1} - \frac{\boldsymbol{I}}{\sigma}\right] (\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H(\boldsymbol{S}\boldsymbol{A}^H\boldsymbol{A} + \sigma\boldsymbol{I})\boldsymbol{A}(\boldsymbol{A}^H\boldsymbol{A})^{-1} \left[(\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{S} + \sigma\boldsymbol{I})^{-1} - \frac{\boldsymbol{I}}{\sigma}\right] \\
&= \frac{1}{\sigma^2}\boldsymbol{S}\boldsymbol{A}^H\boldsymbol{A}(\boldsymbol{S}\boldsymbol{A}^H\boldsymbol{A} + \sigma\boldsymbol{I})^{-1}(\boldsymbol{S}\boldsymbol{A}^H\boldsymbol{A} + \sigma\boldsymbol{I})(\boldsymbol{A}^H\boldsymbol{A})^{-1}(\boldsymbol{S}\boldsymbol{A}^H\boldsymbol{A} + \sigma\boldsymbol{I})^{-1}\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{S} \\
&= \frac{1}{\sigma^2}\boldsymbol{S}(\boldsymbol{S}\boldsymbol{A}^H\boldsymbol{A} + \sigma\boldsymbol{I})^{-1}\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{S} = \frac{1}{\sigma^2}\boldsymbol{S}\boldsymbol{A}^H\boldsymbol{R}^{-1}\boldsymbol{A}\boldsymbol{S},
\end{aligned}$$

where the last equality follows from (24). Also, $\boldsymbol{d}_j{}^H \boldsymbol{P}^{\perp} \boldsymbol{d}_i$ is element (j, i) of $\boldsymbol{D}^H\boldsymbol{P}^{\perp}\boldsymbol{D}$. So we can write

$$\boldsymbol{C} = \frac{2}{\sigma N}\operatorname{Re}\big\{\boldsymbol{D}^H\boldsymbol{P}^{\perp}\boldsymbol{D} \odot \boldsymbol{S}\boldsymbol{A}^H\boldsymbol{R}^{-1}\boldsymbol{A}\boldsymbol{S}\big\}\,.$$

We omit the limit, since $\boldsymbol{C} \to \boldsymbol{0}$ as $N \to \infty$. Now, we get the estimation error covariance matrix as

$$\begin{aligned}
\boldsymbol{C}_{\text{uml}} =& N\left[\frac{2}{\sigma}\operatorname{Re}\big\{\boldsymbol{D}^H\boldsymbol{P}^{\perp}\boldsymbol{D} \odot (\boldsymbol{S}\boldsymbol{A}^H\boldsymbol{R}^{-1}\boldsymbol{A}\boldsymbol{S})^T\big\}\right]^{-1} \frac{2}{\sigma N}\operatorname{Re}\big\{\boldsymbol{D}^H\boldsymbol{P}^{\perp}\boldsymbol{D} \odot \boldsymbol{S}\boldsymbol{A}^H\boldsymbol{R}^{-1}\boldsymbol{A}\boldsymbol{S}\big\} \\
&\cdot \left[\frac{2}{\sigma}\operatorname{Re}\big\{\boldsymbol{D}^H\boldsymbol{P}^{\perp}\boldsymbol{D} \odot (\boldsymbol{S}\boldsymbol{A}^H\boldsymbol{R}^{-1}\boldsymbol{A}\boldsymbol{S})^T\big\}\right]^{-1} \\
=& \frac{\sigma}{2}\left[\operatorname{Re}\big\{\boldsymbol{D}^H\boldsymbol{P}^{\perp}\boldsymbol{D} \odot (\boldsymbol{S}\boldsymbol{A}^H\boldsymbol{R}^{-1}\boldsymbol{A}\boldsymbol{S})^T\big\}\right]^{-1}.
\end{aligned}$$

Hence, we get the Cramér-Rao bound

$$\boldsymbol{B}_u = \frac{\boldsymbol{C}_{\text{uml}}}{N}.$$

IV. Asymptotic Comparison between ML Estimates under CMA and UMA

We show that $\boldsymbol{C}_{\text{cml}} \succeq \boldsymbol{C}_{\text{uml}}$. We have

$$\boldsymbol{C}_{\text{cml}} = \frac{\sigma}{2}[\operatorname{Re}\big\{\boldsymbol{H} \odot \boldsymbol{S}^T\big\}]^{-1}[\operatorname{Re}\big\{\boldsymbol{H} \odot (\boldsymbol{S}\boldsymbol{W}\boldsymbol{S})^T\big\}][\operatorname{Re}\big\{\boldsymbol{H} \odot \boldsymbol{S}^T\big\}]^{-1},$$

where $\boldsymbol{H} \triangleq \boldsymbol{D}^H\boldsymbol{P}^{\perp}\boldsymbol{D}$ and $\boldsymbol{W}^{-1} = \boldsymbol{S}\boldsymbol{A}^H\boldsymbol{R}^{-1}\boldsymbol{A}\boldsymbol{S}$. From [6] we have

$$[\operatorname{Re}\{\boldsymbol{X} \odot \boldsymbol{Z}\}]^{-1}[\operatorname{Re}\{\boldsymbol{X} \odot \boldsymbol{Y}\}][\operatorname{Re}\{\boldsymbol{X} \odot \boldsymbol{Z}\}]^{-1} \geq [\operatorname{Re}\big\{\boldsymbol{X} \odot \boldsymbol{Y}\boldsymbol{Z}^{-1}\boldsymbol{Y}\big\}]^{-1}, \tag{27}$$

where $\boldsymbol{X}$ and $\boldsymbol{Z}$ are positive definite and $\boldsymbol{Y}$ is Hermitian. By using (27) and doing some identification, we can show that $\boldsymbol{C}_{\text{cml}} \succeq \boldsymbol{C}_{\text{uml}}$.

In order to prove (27) one use that [7]

$$\begin{bmatrix} \boldsymbol{X} & \boldsymbol{Y} \\ \boldsymbol{Y}^T & \boldsymbol{Z} \end{bmatrix} \succeq \mathbf{0} \quad \Leftrightarrow \quad \boldsymbol{Z} \succeq \mathbf{0}, \;\; \boldsymbol{X} - \boldsymbol{Y}\boldsymbol{Z}^{-1}\boldsymbol{Y}^T \tag{28}$$

for Hermitian $\boldsymbol{X}$ and $\boldsymbol{Z}$. The left hand side of (28) is a linear matrix inequality (LMI).

REFERENCES

[1] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 10, pp. 1783–1795, Oct. 1990.

[2] P. H. Janssen and P. Stoica, "On the expectation of the product of four matrix-valued gaussian random variables," *IEEE Trans. Automat. Contr.,* vol. 33, pp. 867–870. Sept. 1988.

[3] T. V. K. Chaitanya, "Notes for Asymptotic Analysis Examples," Mar. 2011.

[4] S. K. Mohammed, "Conditional versus unconditional direction of arrival estimation - Key notes," Apr. 2011.

[5] B. Ottersten and L. Ljung, "Asymptotic results for sensor array processing," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing,* Glasgow, Scotland, May 1989, pp. 2266–2269.

[6] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramér-Rao bound: Further Results and Comparisons," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 12, pp. 2140–2150, Dec. 1990.

[7] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, SIAM, 1994.